

A comparison of native and non-native English-speaking groups' understanding of the vocabulary contained within the 16PF (SA92)

Taryn Wallis* and Martin Birt

Organisational Psychology & Human Resource Management, School of Management Studies, University of Cape Town, Private Bag, Rondebosch, 7701
E-mail: twallis@commerce.uct.ac.za

Part of the qualitative phase of Abrahams and Mauer's (1999b) study was replicated at the University of Cape Town with both native English-speaking and non-native English-speaking students. Participants were required to provide synonyms for 135 words contained within the 16PF (SA92) in order to ascertain the extent of problems with the language in the test. When following Abrahams and Mauer's original methodology strictly, results seemed to indicate that both groups struggled with the language contained in the 16PF (SA92). However, less rigid marking, taking into account everyday usage of the words, showed that both groups did understand the words, although they were prevented from displaying this by the original restrictive method. It is suggested that more research be conducted with instruments such as the 16PF in order to obtain a fuller understanding of the extent to which language can affect scores obtained.

*To whom correspondence should be addressed

Marsella, Dubanoski, Hamada and Morse (2000) acknowledge that consensus on a definition of the concept of "culture" is lacking amongst behavioural scientists. Despite this, however, most definitions seem to centre on the aspect of shared patterns of meaning and behaviour that influence the values, attitudes, worldviews and identities of people who belong to the cultural group concerned (Gunsalus & Kelly, 2001; Marsella et al., 2000; McCrae, 2000; Taylor & Boeyens, 1991). This influential aspect of culture on individuals' affective processes has led to many studies and debate on the interrelationship between personality and culture. This research began in the mid-nineteenth century. However, it was only really in the 1950s and 1960s that the issue of how to measure personality cross-culturally came to the fore, with a focus on ensuring that the antecedents, consequences and motivations associated with personality variables were similar across cultures (the problem of conceptual equivalence) (Marsella et al.).

With the awareness of problems of personality measurement across cultures, came the call for personality assessments to be culturally fair (Marsella et al., 2000). Taylor and Boeyens (1991) argued that, despite the structure of objective personality tests often being cited as an advantage to measurement, this structure could also be restrictive by forcing test takers to respond within the narrow confines of the test developer's personal cultural configuration. Since many personality tests still in use in South Africa today were developed and standardised within a Westernised cultural setting, there are implications for the appropriateness of using such instruments with other cultural groups.

While much research has been done to investigate the issue of the cross-cultural comparability of cognitive tests, Taylor and Boeyens (1991) have pointed out that this issue has not received as much attention in the area of personality assessment, although such assessments are widely applied as part of decision-making processes that can impact significantly on people's lives (such as in selection). Some of the reasons put forward by Taylor and Boeyens (1991) for the skewed focus of comparability research include the emotive aspect attached to group differences in performance on cognitive assessments that has driven research in this regard. They state that a tolerant attitude has, however, prevailed in the area of personality assessment, since "somehow it is 'ok' for there to be group differences in personality scores"

(p. 3). This attitude is a result of the complexities inherent in studying the relationship between personality and culture. In fact, Marsella et al. (2000) cite authors (e.g. Shweder, 1991; Sullivan, 1993) who have argued that personality and culture cannot be separated at all. Obviously, this argument only serves to reinforce the point that personality assessment instruments are likely to contain cultural specificities.

Taylor and Boeyens (1991) listed two main types of comparability that should be established when investigating the cross-cultural comparability of personality instruments. The first is that of *construct comparability*. This aims to determine if a particular personality construct actually exists within the cultural groups under study, and also if a particular personality test itself is able to assess this construct in each of the cultural groups. Construct comparability has also been discussed by Fouad (1993) as metric equivalence, which refers to whether the psychometric scale being utilised is measuring the same construct in different cultures.

The second type of comparability is *score comparability*, which aims to determine whether the scores achieved on a personality test are comparable across the different cultural groups (Taylor & Boeyens, 1991). Score comparability can only be determined after it has been shown that construct comparability exists. Fouad (1993) also examines three other types of equivalence: functional equivalence (examining the role or function of specific behaviours in various cultures to ensure that these are viewed in a similar manner if they are to be included in a personality test), conceptual equivalence (ensuring that the concepts to be used are similar in terms of the meaning attached to them by different cultural groups), and linguistic equivalence (ensuring that translated instruments are equivalent to the original).

The issue of language has received much attention in psychological assessment, as this is closely linked to the culture in which the test is developed, and language gives expression to the cultural concepts and constructs that need to be measured (McCrae, 2000; Saunders, 2000). This is a complex area of study, however, as not only are there inter-cultural differences in language usage, but language also evolves and changes over time within cultural groupings.

The field of psychological assessment faces two challenges in this regard. One is the challenge of translating tests, with its associated problem of attaining linguistic equivalence, as discussed by Fouad (1993). Saunders (2000) has acknowledged that many of the attempts to translate tests have resulted in inaccuracies and the distortion of the meaning of items, making them difficult to understand. To illustrate this, she cited a personal experience of providing a Zulu translation of a particular test to her Zulu-speaking clients, only to have them reject this in favour of the original English version, as the latter was more understandable.

The second challenge concerns the assessment of individuals in a language that is not their vernacular. Since test translation is a time-consuming and expensive process, many tests remain in their original language. South Africa is a multi-cultural and multi-linguistic country. However, many assessment instruments in current usage are in the English language. Saunders (2000) has pointed out that the assumption is often made that someone who is able to speak English fairly proficiently will also be able to read and write well enough to be able to cope with the level of English required by an assessment instrument. This assumption is, however, unfounded. Ellis and Mead (2000) have also sounded a caution that language differences can sometimes be quite subtle and therefore go undetected. A lack of proficiency in English, therefore, may result in inaccurate test results being obtained.

In the South African context, the language spoken by a particular sub-culture is also often linked to its Apartheid era population group classification (Black, White, Coloured and Indian). Although it is acknowledged that there are different language groupings within each population category as well as some language groupings that transcend these population categories, it also needs to be acknowledged that the majority of Black South Africans speak English as either a second or third language. The potential for unfair discrimination against this group therefore cannot be ruled out when English assessment instruments are used for purposes such as selection. Although the use of such tests is regulated by the *Employment Equity Act* (1998), and responsible test users should never base a decision solely on the information obtained from one questionnaire, this potential for discrimination has been much reported by the media, with a consequent negative perception of such tests being held by both employees and employers (Sehlapelo & Terre Blanche, 1996). Employees, particularly, have been reported to see the tests as ineffective in the measurement of valid individual differences and often cannot see the relationship between personality test items and the job skills required in their everyday work (Sehlapelo & Terre Blanche, 1996). It is therefore vital to conduct research on the comparability of test scores across cultures, as well as into the reasons for any differences obtained, in order to prevent what Shuttleworth-Jordan (1996) has termed “an attitude of nihilism with respect to test usage” (p. 96) arising.

Two South African studies that have addressed the issue of culture and language with respect to personality assessment are Abrahams and Mauer’s (1999a) study on the comparability of the constructs of the Sixteen Personality Factor Questionnaire (16PF) (Cattell, Eber & Tatsuoka, 1992, cited in Abrahams & Mauer, 1999a) and Abrahams and Mauer’s (1999b) study on the qualitative and statistical impact of home language on responses to the items of the 16PF (cf. Kerka, 1995). Both of these studies utilised the SA92 version of the 16PF, a version developed with the intention of ensuring that the test was appropriate for all cultural groups in South Africa (Abrahams & Mauer, 1999a).

Abrahams and Mauer’s (1999a) study showed significant differences between race groups with regard to factor structures as well as response patterns. In a subsequent qualitative phase (Abrahams & Mauer, 1999b), 71 second-year Industrial Psychology students, none of whom spoke English as a first language, struggled to provide synonyms for 136 words contained in the 16PF, and 10 Honours students (also non-native English speakers) experienced problems understanding the sentence construction of some of the items. These results were taken to indicate that the differences in endorsement rates between races on the 16PF could, in part, be attributed to language difficulties. Other differences that could not be explained by language were attributed to “cultural factors”. In Taylor and Boeyens’ (1991) research, cultural factors were specified as cultural mores (cultural differences in acceptability of types of behaviour), cultural beliefs, situational and experiential factors (differences in the day-to-day experiences of Black and White South Africans), and the social desirability attached to particular activities in different cultures.

Although Abrahams and Mauer (1999a) certainly found score differences between race groups on the 16PF that are indicative of potential cultural bias, the fact that Abrahams and Mauer (1999b) did not compare their sample of students with a first language English-speaking group does not allow for valid conclusions to be drawn about the true extent to which language difficulties can be said to be responsible for these score differences. If language difficulties are at the heart of score discrepancies, there should be significant differences between the abilities of first language and second or third language English speakers to provide correct synonyms for the words contained in the 16PF. It is for this reason that this aspect of Abrahams and Mauer’s (1999b) study was replicated, using these two groups in order to compare their results and thus to contribute to the debate on the issue of language usage in the 16PF as well as on personality testing in general. To maintain consistency with Abrahams and Mauer’s (1999b) terminology, first language English speakers will be referred to as “native English speakers” and those students who speak English as either a second or third language will be referred to as “non-native English speakers”.

METHOD

Participants

The sample of participants in this study consisted of 131 second-year Industrial Psychology students at the University of Cape Town. The participants were asked to provide their home language and race. This information is summarised in Table 1

Table 1. Breakdown of sample according to language and race

Race	Native English speakers	Non-native English speakers
White	50	4
Black	4	27
Coloured	27	2
Indian	9	1
Unspecified	6	1
Total	96	35

In terms of gender, 101 (77%) of the participants were female and 30 (23%) were male. No details of gender breakdown were provided by Abrahams and Mauer (1999b). Thus, it is not known how the gender distribution of this sample compares to the previous study.

Procedure

This study aimed to duplicate the procedure followed by Abrahams and Mauer (1999b) with respect to determining the

ability of students to provide synonyms for words encountered in the 16PF.

The students were provided with the same list of words utilized by Abrahams and Mauer (1999b). This list consisted of 136 nouns and adjectives that are found in the 16PF. It was subsequently noticed that Abrahams and Mauer’s list repeated the word “depressed”. Marking therefore took this into account and results are reported in terms of 135 words. Students were instructed to provide synonyms for the words contained in the list. This exercise was completed during one of the double period tutorials available for this group of students, the same time allocation provided to students in Abrahams and Mauer’s (1999b) study.

Data analysis

Synonyms were marked as correct or incorrect on the basis of whether they appeared in the Pocket Oxford Dictionary (Oxford University Press, 1992), Merriam-Webster’s Collegiate Dictionary (Merriam-Webster, 2001) and the Ultralingua English Dictionary (Ultra Lingua Net, 2001). Efforts were made to source the same dictionaries used by Abrahams and Mauer (1999b), but these were unsuccessful. The above-mentioned dictionaries were therefore chosen because they are comprehensive and, in the case of Merriam-Webster’s Collegiate Dictionary (on-line), provide a thesaurus service as well (Merriam-Webster’s Collegiate Thesaurus, which can be consulted concurrently with the dictionary entries).

In order to ensure that the results were accurate, the marking was conducted by one researcher and checked by the second researcher.

RESULTS

When following Abrahams and Mauer’s (1999b) method strictly, the results of this study indicated the same patterns for both the native and non-native English-speaking groups. Neither group was able to provide acceptable synonyms for the words contained in the list most of the time. These results are reported in Appendices A and B for each group under the first two columns (“Strict (n)” and “%”). These two columns show, respectively, the number of students responding correctly when their responses were marked strictly according to the dictionaries and the corresponding percentage of correct responses.

Table 2 summarises the results by providing the number of words for which less than 50% of the sample could provide acceptable synonyms, that is, the number of words answered incorrectly by at least half of the sample.

Table 2. Number of words for which incorrect synonyms provided by at least 50% of sample (strict marking)

	Number of words
Abrahams & Mauer’s (1999b) original sample (non-native English speakers)	78
Current study (non-native English speakers)	113
Current study (native English speakers)	105

When marking the results, however, it was noticed that many students provided synonyms that were not strictly correct according to the dictionaries, but which would be understood as correct in everyday usage, and which indicated that the students did know what the words in the list meant. These included Afrikaans translations (e.g. confused = *deurmekaar*), subtle synonyms (e.g. temptation = chocolate), and words used in

everyday English as meaning the same thing although they are technically different (e.g. beach = sea). To ensure that this less rigid marking was not biased by the one researcher’s subjective views, the second researcher checked the marking of the other researcher. The results therefore reflect both researchers’ understanding of the colloquial usage of language. These results are also reported in the Appendices in the columns headed “Relaxed(n)” (the number of additional students responding correctly when less rigid marking was applied), “S+R(n)” (the number of correct responses from both the strictly marked and relaxed marking groups), and “%” (the corresponding percentage of correct responses for both the strict and relaxed marking). The changed figures are summarised in Table 3

Table 3. Number of words for which incorrect synonyms provided by at least 50% of sample (relaxed marking)

	Number of words
Abrahams & Mauer’s (1999b) original sample (non-native English speakers)	N/A
Current study (non-native English speakers)	7
Current study (native English speakers)	0

DISCUSSION

In the current study, when following Abrahams and Mauer’s (1999b) method strictly, the results obtained for non-native English-speakers were similar to those obtained by Abrahams and Mauer’s (1999b) sample. Participants were unable to provide the correct synonyms for the words most of the time. However, utilising a comparative approach has provided a significant and surprising finding, in that the results indicate the same patterns for both the native English-speaking and non-native English-speaking groups in the current study. In other words, when using this methodology, neither group could provide acceptable synonyms for the words most of the time. The fact that both groups in the study got more synonyms incorrect than Abrahams and Mauer’s sample could possibly be explained by the use of different dictionaries. However, what is important to note is that, within the current study, both the native English-speaking and non-native English-speaking students were compared to the same standard, with very similar results. This suggests that either the language in the 16PF is too difficult for people of all language groups to understand, or that problems with the original methodology produced a skewed result.

The inclusion of less rigid marking seems to indicate that the problem may be more methodological than language-related, as the less rigid marking shows that both groups did understand the words contained in the list. These results also seem to corroborate comments made by the students in a debriefing after the exercise. Both the native and non-native English-speaking students indicated that they found it difficult to think of synonyms for the words, although they understood perfectly well what the words meant and would have been able to use them in a sentence. One non-native English-speaking student wrote the following comment on her questionnaire: “If you want to see whether people understand what the words mean, asking for synonyms is probably the easiest, but not the best option. It’s easier to tell in a sentence what a word means, and besides, one word rarely reflects the true meaning of another.” The non-native English-speaking students, in particular, had an issue with the methodology, as they claimed that it discriminated against them. This was because providing synonyms required them to first translate the word into their own language, then find a synonym for the word in their own language, and then translate that synonym back into English, even though they understood

what the word meant. Saunders (2000) has argued that this type of internal translation procedure is common in second language English-speakers when being assessed. For personality tests, in which no time limits are imposed, this may add to the time taken, but will not necessarily disadvantage the person in terms of results. However, this translation procedure could certainly be detrimental to the results of non-native English speakers when undergoing timed cognitive assessments.

Although the results from the current study seem to indicate that the issue of language is perhaps not as serious as previously thought, it must be noted that these results do not necessarily rule out language problems as the reason for differences found by Abrahams and Mauer (1999a) between students of different population groups on the 16PF. Indeed, given the fact that, similarly to Abrahams and Mauer's (1999b) original study, nine native English-speaking and two non-native English-speaking students from the current sample thought that a "bookkeeper" was a "librarian", it seems likely that there are certain words in the 16PF that may not be understood by any of the language groups. This may be less of a cultural issue though, and more related to changes in language usage over time. In a debriefing after the exercise, students seemed unfamiliar with the concept of "bookkeeper", but familiar with the term, "accountant". Language changes will need to be made in order to keep the 16PF as up-to-date as possible, and the language used in new versions of the test, such as the 15FQ+ (previously the 16PF Industrial) (Psytech International, 2000) that is also currently in usage, should be checked.

The fact remains that Abrahams and Mauer (1999a) did find differences in scores as well as in the factor structures between students of different population groups on the 16PF. This is indicative of a potential culture bias in the instrument, a possibility every time a language-based assessment instrument is used (Saunders, 2000). Since personality tests tend to be quite reliant on language, they are particularly susceptible to such culture bias.

However, Saunders (2000) cautions against such score differences always being interpreted as evidence of culture bias, as the possibility cannot be ruled out conclusively that score differences are reflecting real personality differences between people, and in fact may point to areas in which development may be necessary. In highlighting this problem, Saunders (2000) has touched on the area that is at the heart of the complexities in the debate on cultural influences on personality. The results of the current study as well as the fact that Abrahams and Mauer (1999b) found differences that could not be explained in terms of language problems, suggest that the underlying cause of score differences may in fact be attributable to deeper cultural factors. This is an area of research that is, as Abrahams and Mauer (1999b) acknowledged, "bound to set exacting demands on, and unusual challenges for, researchers" (p. 76) because of the difficulties involved in determining exactly what aspects of a culture are affecting test scores. It is, however, an area that must be addressed, as merely changing the language of questionnaires such as the 16PF or developing a new questionnaire may not solve the problem. As the method employed by Abrahams and Mauer (1999b) to determine that score differences were to a large extent due to language barriers seems to have been ineffective, it is recommended that studies utilising different methodologies be undertaken to ascertain the extent to which language barriers may truly be responsible for differences.

A recent study by Gunsalus and Kelly (2001) seems to support the idea that score differences are attributable to cultural factors

deeper than language. They investigated the effect of the Korean culture on the Millon Clinical Multiaxial Inventory-III (MCMI-III) (Millon, 1997, cited in Gunsalus & Kelly, 2001) by comparing the personality profiles of Korean and American college students. In their study, an effort was made to control for language differences by choosing Korean students who were majoring in either English or English Education from a Korean university known for the high quality of its English training. Gunsalus and Kelly's findings indicated that the Korean and American groups differed significantly on seven of the eleven scales measured by the instrument. Although the possibility cannot be ruled out that these results reflect language problems (despite an effort to control this variable), the results were taken to show that significant differences between the two cultures were reflected in responses to the MCMI-III (Gunsalus & Kelly, 2001). These differences were generally consistent with previous research that suggested that the Korean culture is other-oriented and "passive", compared to the more individualistic and "active" American culture (e.g. Liem & Lim, 2000, cited in Gunsalus & Kelly, 2001).

Research conducted in South Africa recently by Heuchert, Parker, Stumpf and Myburgh (2000) with the Revised NEO Personality Inventory (NEO-PI-R) (Costa & McCrae, 1992, cited in Heuchert *et al.*, 2000) showed that the structure of the five-factor model was reproduced for the entire sample as well as for the Black and White sub-groups. Despite the structure of personality being similar across the racial groups, however, they found significant differences in the mean scores of some domains and facets. These were attributed to social, economic and cultural differences between the races. It must be noted, however, that the size of the Black sub-group in Heuchert *et al.*'s (2000) study was smaller than that recommended for factor analysis, and the authors therefore cautioned that results should be treated as exploratory.

The problem of matching factor structures, as experienced by Heuchert *et al.* (2000), is one that has beset studies in the area of culture and personality. Abrahams and Mauer (1999a) were unable to replicate Cattell's original factor structure across the race groups in their study. Marsella *et al.* (2000) have, however, cautioned that matched factor structures do not necessarily automatically translate into equivalence across cultures, citing reasons such as factors accounting for different proportions of variance across cultures, items within factors loading differently, and differences in means and variances being present.

The question then arises as to whether different norm groups should be developed for people of different cultural groups. Both Austin (1999) and Saunders (2000) have commented that this can be useful in areas such as career assessment and selection. For selection purposes, Saunders (2000) recommended making use of two norm groups when assessing individuals who may be disadvantaged in terms of language barriers, namely, a norm from a White group and one from a Black group, in order to ascertain the magnitude of score differences that could be attributable to culture or language factors.

The use of separate norm groups, however, may be complicated by the problem of ethnocultural identity – the extent to which an individual accepts and displays the traditions and practices of a particular cultural group (Marsella *et al.*, 2000). The assumption that someone from a specific ethnic group identifies with that group's culture may actually be invalid. In fact, Fouad (1993) has argued that descriptions of cultures tend to be general and conceal within-group variance. In order to determine the extent to which someone has adopted the values of a particular culture,

Foad (1993) has recommended an examination of people's levels of acculturation, assimilation, socio-economic status, family history and individual history. In this regard, Marsella *et al.* (2000) advised that university students should not be used as samples in cross-cultural studies of personality, as students are more exposed to, and thus more likely to adopt, aspects of Western culture. A measure of ethnocultural identity should therefore be a necessary component in research on personality and culture.

Lastly, another challenge to the study of personality and culture is the criticism that the concept of personality itself is based on Western ideals of individualism and a bounded self, and may not even be meaningful or relevant in non-Western cultures (Marsella *et al.*, 2000).

CONCLUSION

A plausible argument put forward in an effort to explain the relationship between personality and culture suggests that certain traits, such as the Big Five, may be universal tendencies, but that cultural factors determine the manner in which these are expressed, the situations in which they are acceptable, and the meanings or value judgements assigned to them (Marsella *et al.*, 2000; McCrae, 2000).

This argument highlights generally accepted tenets regarding personality and culture, namely, that personality is influenced by cultural experiences and that the structure of, and processes within, personality can vary widely between considerably different cultures (Marsella *et al.*, 2000). Because of this, assessment practitioners need to exercise caution in the construction and use of personality instruments, and should possess an understanding of the culture within which the instrument will be used. These caveats have been expressed in South Africa, with critics of instruments such as the 16PF arguing that it should not be used unconditionally with the multi-cultural workforce of this country (Abrahams & Mauer, 1999b). Current responsible test users in this country would subscribe to this view, using assessment instruments supported by research where possible, never using any instrument in isolation, and making use of an in-depth feedback interview to increase the accuracy of results. At the same time, more research must be done on the effect of cultural factors and language barriers on scores on personality instruments such as the 16PF, in order to ascertain the extent of the influence of these factors and barriers. Disregarding a particular test, or testing in general, should not be seen as an option, as this may not provide a solution.

REFERENCES

Abrahams, F. & Mauer, K. F. (1999a). The comparability of the constructs of the 16PF in the South African context. *Journal of Industrial Psychology*, 25(1), 53 – 59.

Abrahams, F. & Mauer, K. F. (1999b). Qualitative and statistical impacts of home language on responses to the items of the Sixteen Personality Factor Questionnaire (16PF) in South Africa. *South African Journal of Psychology*, 29(2), 76 – 86.

Austin, J. T. (1999). Culturally sensitive career assessment: A quandary. *ERIC Digest*, 210.

http://www.ed.gov/databases/ERIC_Digests/ed434246.html. Accessed on 14 November 2002.

Ellis, B. B. & Mead, A. D. (2000). Assessment of the measurement equivalence of a Spanish translation of the 16PF questionnaire. *Educational & Psychological Measurement*, 60(5), 787–808.

Employment Equity Act (1998). *Employment Equity Act, No. 55 of 1998*. Republic of South Africa. <http://www.labour.gov.za/docs/legislation/eea/act98-055.html>. Accessed on 28 July 2003.

Fouad, N. A. (1993). Cross-cultural vocational assessment. *Career Development Quarterly*, 42(1), 4–14.

Gunsales, A. C. & Kelly, K. R. (2001). Korean cultural influences on the Millon Clinical Multiaxial Inventory III. *Journal of Mental Health Counseling*. <http://www.epnet.com/ehost/login/html>. Accessed on 30 November 2001.

Heuchert, J. W. P., Parker, W. D., Stumpf, H. & Myburgh, C. P. H. (2000). The five-factor model of personality in South African college students. *American Behavioral Scientist*. <http://www.epnet.com/ehost/login/html>. Accessed on 30 November 2001.

Kerka, S. (1995). Adult career counseling in a new age. *ERIC Digest*, 167. <http://ericacve.org/docs/adult.html>. Accessed on 14 November 2002.

Marsella, A. J., Dubanoski, J., Hamada, W. C. & Morse, H. (2000). The measurement of personality across cultures. *American Behavioral Scientist*. <http://www.epnet.com/ehost/login/html>. Accessed on 30 November 2001.

McCrae, R. R. (2000). Trait psychology and the revival of personality and culture studies. *American Behavioral Scientist*. <http://www.epnet.com/ehost/login/html>. Accessed on 30 November 2001.

Merriam-Webster (2001). *Collegiate dictionary*. <http://www.m-w.com>. Accessed on 30 November 2001.

Oxford University Press (1992). *The pocket Oxford dictionary*. New York: Oxford University Press.

Psytech International (2000). *The 16PF Industrial technical manual*. Bedfordshire: Psytech International.

Saunders, E. (2000). *Assessing human competence: Practical guidelines for the South African manager*. Randburg: Knowledge Resources.

Sehlapelo, M. & Terre Blanche, M. (1996). Psychometric testing in South Africa: Views from above and below. *Psychology in Society*, 21, 49 – 59.

Shuttleworth-Jordan, A. B. (1996). On not reinventing the wheel: A clinical perspective on culturally relevant test usage in South Africa. *South African Journal of Psychology*. <http://www.epnet.com/ehost/login.html>. Accessed on 15 September 1999.

Taylor, T. R. & Boeyens, J. C. A. (1991). The comparability of the scores of blacks and whites on the South African Personality Questionnaire: An exploratory study. *South African Journal of Psychology*, 21 (1), 1–11.

Ultra Lingua Net (2001). *Ultralingua English dictionary*. <http://www.ultralingua.net>. Accessed on 30 November 2001.

APPENDIX A

NATIVE ENGLISH-SPEAKING / TOTAL (N = 96)

	Strict (n)	%	Relaxed (n)	S + R (n)	%
Absentminded	63	65.6	17	80	83.3
Accuracy	53	55.2	31	84	87.5
Activities	4	4.2	78	82	85.4
Admire	45	46.9	43	88	91.7
Afraid	17	17.7	79	96	100.0
Aid	80	83.3	14	94	97.9
Appliances	51	53.1	36	87	90.6
Argument	49	51.0	43	92	95.8
Artistic	0	0.0	90	90	93.8
Assistants	87	90.6	8	95	99.0
Attention	17	17.7	53	70	72.9
Attractive	26	27.1	66	92	95.8
Avoid	16	16.7	59	75	78.1
Background	9	9.4	77	86	89.6
Basic	38	39.6	51	89	92.7
Battles	50	52.1	44	94	97.9
Beach	16	16.7	70	86	89.6
Beauty	14	14.6	63	77	80.2
Betrays	10	10.4	69	79	82.3
Bookkeeper	0	0.0	76	76	79.2
Bossy	28	29.2	56	84	87.5
Brag	57	59.4	34	91	94.8
Calculated	3	3.1	81	84	87.5
Calm	27	28.1	65	92	95.8
Candle	2	2.1	75	77	80.2
Care	38	39.6	42	80	83.3
Career	19	19.8	72	91	94.8
Challenge	8	8.3	63	71	74.0
Characteristics	60	62.5	23	83	86.5
Children	76	79.2	17	93	96.9
Citizens	50	52.1	36	86	89.6
Clumsy	1	1.0	52	53	55.2
Committee	37	38.5	43	80	83.3
Composed	36	37.5	48	84	87.5
Concerned	82	85.4	5	87	90.6
Confused	12	12.5	68	80	83.3
Consequences	58	60.4	26	84	87.5
Convenient	6	6.3	75	81	84.4
Conversations	52	54.2	38	90	93.8
Coordination	3	3.1	55	58	60.4
Criticism	7	7.3	51	58	60.4
Danger	4	4.2	65	69	71.9
Depressed	72	75.0	19	91	94.8
Diligence	33	34.4	32	65	67.7
Discouraged	8	8.3	60	68	70.8
Dishonest	16	16.7	73	89	92.7
Disloyal	13	13.5	60	73	76.0
Downhearted	1	1.0	78	79	82.3
Dreamer	11	11.5	65	76	79.2
Efficient	16	16.7	59	75	78.1
Electrical	1	1.0	57	58	60.4
Embarrassed	35	36.5	31	66	68.8
Emotional	2	2.1	73	75	78.1
Entertaining	15	15.6	65	80	83.3
Enthusiastic	35	36.5	52	87	90.6
Excitement	1	1.0	77	78	81.3
Exercise	40	41.7	47	87	90.6
Factory	5	5.2	80	85	88.5
Flame	0	0.0	86	86	89.6
Forgive	5	5.2	58	63	65.6
Functions	17	17.7	63	80	83.3
Gathering	20	20.8	72	92	95.8
Guests	46	47.9	33	79	82.3
Happy-go-lucky	29	30.2	43	72	75.0

Headline	7	7.3	72	79	82.3
Honesty	33	34.4	50	83	86.5
Imagination	20	20.8	63	83	86.5
Immediately	76	79.2	16	92	95.8
Impractical	5	5.2	66	71	74.0
Independent	11	11.5	70	81	84.4
Influence	18	18.8	59	77	80.2
Intellectual	0	0.0	91	91	94.8
Interesting	4	4.2	61	65	67.7
Interruptions	13	13.5	66	79	82.3
Invention	20	20.8	58	78	81.3
Jealousy	3	3.1	66	69	71.9
Level-headed	4	4.2	77	81	84.4
Logical	0	0.0	79	79	82.3
Love story	0	0.0	85	85	88.5
Machines	1	1.0	79	80	83.3
Manners	13	13.5	64	77	80.2
Military	72	75.0	10	82	85.4
Mishaps	19	19.8	66	85	88.5
Modern	25	26.0	58	83	86.5
Neighbours	19	19.8	56	75	78.1
Nerves	0	0.0	52	52	54.2
Obedying	13	13.5	68	81	84.4
Occasionally	29	30.2	62	91	94.8
Opinion	25	26.0	63	88	91.7
Opportunities	48	50.0	24	72	75.0
Outgoing	49	51.0	39	88	91.7
Overexcited	0	0.0	78	78	81.3
Peculiar	69	71.9	21	90	93.8
Persuade	38	39.6	47	85	88.5
Photographic	0	0.0	77	77	80.2
Queue	82	85.4	13	95	99.0
Reaction	45	46.9	32	77	80.2
Rejected	10	10.4	75	85	88.5
Repairing	89	92.7	3	92	95.8
Reporter	0	0.0	88	88	91.7
Routine	1	1.0	80	81	84.4
Scent	91	94.8	2	93	96.9
Scientist	1	1.0	67	68	70.8
Self-centred	33	34.4	62	95	99.0
Sensitive	9	9.4	74	83	86.5
Setbacks	25	26.0	46	71	74.0
Spirited	4	4.2	71	75	78.1
Social	4	4.2	75	79	82.3
Statue	16	16.7	60	76	79.2
Stranger	34	35.4	42	76	79.2
Strict	13	13.5	69	82	85.4
Superior	19	19.8	68	87	90.6
Talent	56	58.3	32	88	91.7
Temptations	0	0.0	51	51	53.1
Thorn	0	0.0	70	70	72.9
Accept	51	53.1	28	79	82.3
Cheerful	85	88.5	8	93	96.9
Familiar	17	17.7	62	79	82.3
Firm	25	26.0	69	94	97.9
Hire	47	49.0	42	89	92.7
Hotel	9	9.4	73	82	85.4
Mechanical	8	8.3	58	66	68.8
Organised	4	4.2	70	74	77.1
Serious	5	5.2	52	57	59.4
Seaside	15	15.6	75	90	93.8
Sufficient	79	82.3	4	83	86.5
Abroad	79	82.3	10	89	92.7
Ability	47	49.0	38	85	88.5
Court	7	7.3	68	75	78.1
Explore	20	20.8	68	88	91.7
Nasty	22	22.9	71	93	96.9
Salaries	61	63.5	32	93	96.9
Sheltered	35	36.5	52	87	90.6

APPENDIX B

NON-NATIVE ENGLISH-SPEAKING / TOTAL (N = 35)

	Strict (n)	%	Relaxed (n)	S + R (n)	%
Absentminded	10	28.6	17	27	77.1
Accuracy	14	40.0	16	30	85.7
Activities	2	5.7	23	25	71.4
Admire	4	11.4	26	30	85.7
Afraid	7	20.0	26	33	94.3
Aid	33	94.3	0	33	94.3
Appliances	13	37.1	13	26	74.3
Argument	22	62.9	7	29	82.9
Artistic	0	0.0	27	27	77.1
Assistants	29	82.9	4	33	94.3
Attention	6	17.1	15	21	60.0
Attractive	7	20.0	26	33	94.3
Avoid	7	20.0	19	26	74.3
Background	2	5.7	26	28	80.0
Basic	10	28.6	21	31	88.6
Battles	22	62.9	11	33	94.3
Beach	7	20.0	18	25	71.4
Beauty	1	2.9	23	24	68.6
Betrays	4	11.4	25	29	82.9
Bookkeeper	0	0.0	27	27	77.1
Bossy	10	28.6	20	30	85.7
Brag	10	28.6	18	28	80.0
Calculated	3	8.6	26	29	82.9
Calm	11	31.4	20	31	88.6
Candle	0	0.0	22	22	62.9
Care	15	42.9	12	27	77.1
Career	10	28.6	20	30	85.7
Challenge	2	5.7	21	23	65.7
Characteristics	16	45.7	13	29	82.9
Children	31	88.6	4	35	100.0
Citizens	18	51.4	12	30	85.7
Clumsy	1	2.9	20	21	60.0
Committee	13	37.1	18	31	88.6
Composed	8	22.9	22	30	85.7
Concerned	20	57.1	3	23	65.7
Confused	2	5.7	22	24	68.6
Consequences	20	57.1	11	31	88.6
Convenient	1	2.9	22	23	65.7
Conversations	14	40.0	13	27	77.1
Coordination	2	5.7	22	24	68.6
Criticism	0	0.0	20	20	57.1
Danger	6	17.1	19	25	71.4
Depressed	14	40.0	9	23	65.7
Diligence	12	34.3	6	18	51.4
Discouraged	2	5.7	22	24	68.6
Dishonest	7	20.0	25	32	91.4
Disloyal	6	17.1	21	27	77.1
Downhearted	0	0.0	21	21	60.0
Dreamer	6	17.1	18	24	68.6
Efficient	2	5.7	20	22	62.9
Electrical	1	2.9	16	17	48.6
Embarrassed	12	34.3	9	21	60.0
Emotional	0	0.0	25	25	71.4
Entertaining	3	8.6	20	23	65.7
Enthusiastic	9	25.7	21	30	85.7
Excitement	1	2.9	22	23	65.7
Exercise	6	17.1	20	26	74.3
Factory	2	5.7	27	29	82.9
Flame	0	0.0	25	25	71.4
Forgive	3	8.6	14	17	48.6
Functions	6	17.1	17	23	65.7
Gathering	6	17.1	26	32	91.4
Guests	25	71.4	6	31	88.6
Happy-go-lucky	2	5.7	16	18	51.4

Headline	0	0.0	23	23	65.7
Honesty	5	14.3	24	29	82.9
Imagination	4	11.4	18	22	62.9
Immediately	25	71.4	7	32	91.4
Impractical	0	0.0	27	27	77.1
Independent	4	11.4	24	28	80.0
Influence	5	14.3	20	25	71.4
Intellectual	0	0.0	28	28	80.0
Interesting	0	0.0	19	19	54.3
Interruptions	2	5.7	24	26	74.3
Invention	7	20.0	18	25	71.4
Jealousy	1	2.9	16	17	48.6
Level-headed	1	2.9	18	19	54.3
Logical	0	0.0	26	26	74.3
Love story	0	0.0	21	21	60.0
Machines	2	5.7	23	25	71.4
Manners	8	22.9	15	23	65.7
Military	19	54.3	9	28	80.0
Mishaps	6	17.1	22	28	80.0
Modern	10	28.6	19	29	82.9
Neighbours	3	8.6	18	21	60.0
Nerves	0	0.0	19	19	54.3
Obedying	3	8.6	19	22	62.9
Occasionally	9	25.7	16	25	71.4
Opinion	9	25.7	20	29	82.9
Opportunities	25	71.4	4	29	82.9
Outgoing	14	40.0	12	26	74.3
Overexcited	0	0.0	22	22	62.9
Peculiar	19	54.3	7	26	74.3
Persuade	19	54.3	10	29	82.9
Photographic	1	2.9	16	17	48.6
Queue	23	65.7	6	29	82.9
Reaction	13	37.1	9	22	62.9
Rejected	8	22.9	17	25	71.4
Repairing	24	68.6	3	27	77.1
Reporter	0	0.0	23	23	65.7
Routine	0	0.0	25	25	71.4
Scent	25	71.4	3	28	80.0
Scientist	0	0.0	17	17	48.6
Self-centred	16	45.7	13	29	82.9
Sensitive	2	5.7	19	21	60.0
Setbacks	10	28.6	10	20	57.1
Spirited	2	5.7	18	20	57.1
Social	1	2.9	21	22	62.9
Statue	2	5.7	18	20	57.1
Stranger	10	28.6	13	23	65.7
Strict	2	5.7	20	22	62.9
Superior	1	2.9	25	26	74.3
Talent	13	37.1	12	25	71.4
Temptations	0	0.0	9	9	25.7
Thorn	0	0.0	16	16	45.7
Accept	13	37.1	9	22	62.9
Cheerful	23	65.7	5	28	80.0
Familiar	2	5.7	23	25	71.4
Firm	4	11.4	21	25	71.4
Hire	11	31.4	11	22	62.9
Hotel	3	8.6	18	21	60.0
Mechanical	2	5.7	17	19	54.3
Organised	2	5.7	18	20	57.1
Serious	3	8.6	9	12	34.3
Seaside	5	14.3	20	25	71.4
Sufficient	24	68.6	1	25	71.4
Abroad	22	62.9	4	26	74.3
Ability	8	22.9	20	28	80.0
Court	0	0.0	19	19	54.3
Explore	3	8.6	23	26	74.3
Nasty	3	8.6	22	25	71.4
Salaries	11	31.4	18	29	82.9
Sheltered	4	11.4	24	28	80.0